PORTFOLIO STRUCTURING USING
LOW-DISCREPANCY DETERMINISTIC SEQUENCES

5   This application is a continuation-in-part of pending application No. 08/902,921, *still pending,* which was filed July 30, 1997 as a continuation application of now-abandoned application No. 08/285,902 filed August 4, 1994.

Background of the Invention

      The invention relates to portfolio structuring and, more particularly, to structuring of a portfolio including assets/liabilities such as real estate, stocks, bonds, futures, options, collateralized mortgage obligations, single-premium deferred annuities and insurance contracts.

      Portfolio structuring includes transactions such as buying, holding and selling of assets/liabilities, as well the adoption, continuation, modification or termination of transaction strategies. Typically, decisions concerning transactions and strategies are based on the evaluation of one or several complicated functions of a large number of variables such as a multi-dimensional integral or the inverse of a distribution function, for example.

      In financial securities transactions, which includes the initial sale, the value of a security may be estimated, e.g., based on expected future cash flow. Such cash flow

may depend on variable interest rates, for example, and these and other relevant variables may be viewed as stochastic variables.

5   It is well known that the value of a financial security which depends on stochastic variables can be estimated in terms of a multi-dimensional integral. The dimension of such an integral may be very high.

In collateralized mortgage obligations (CMO), for example, instruments or securities variously called
10   tranches, shares, participations, classes or contracts have cash flows which are determined by dividing and distributing the cash flow of an underlying collection or pool of mortgages on a monthly basis according to pre-specified rules. The present value of a tranche can be estimated on
15   the basis of the expected monthly cash flows over the remaining term of the tranche, and an estimate of the present value of a tranche can be represented as a multi-dimensional integral whose dimension is the number of payment periods of the tranche. For a typical instrument
20   with a 30-year term and with monthly payments, this dimension is 360.

Usually, such a high-dimensional integral can be evaluated only approximately, by numerical integration. This involves the generation of points in the domain of
25   integration, evaluating or "sampling" the integrand at the generated points, and combining the resulting integrand values, e.g., by averaging. Well known for numerical integration in securities valuation is the so-called Monte

Carlo method in which points in the domain of integration are generated at random.

With integrands arising in financial securities valuation, the computational work in combining the sampled

5 values is negligible as compared with producing the integrand values. Thus, numerical integration methods in securities valuation may be compared based on the number of samples required for obtaining a sufficiently accurate approximation to the integral.

10 Sampling techniques are useful also as applied to formulations which need not be in terms of an integral. In particular, in portfolio structuring, this applies to a quantity known as value at risk (VAR) which may be defined as the worst potential loss of value of a portfolio over a

15 period of time under normal market conditions, at a specified confidence level. While value at risk can be expressed as a definite limit in a one-sided indefinite integral (see Philippe Jorion, Value at Risk, Irwin Professional Publishing, 1997, p. 88), a more practical

20 formulation is in terms of the inverse of a cumulative distribution function.

Summary Of The Invention

A preferred method for estimating the value of

25 a financial security involves numerical integration unlike Monte Carlo integration in that an integrand is sampled at deterministic points having a low-discrepancy property. As compared with the Monte Carlo method, significant advantages

are realized with respect to speed, accuracy, and dependability.

A computed estimate can be used in setting an offering price, e.g., somewhat higher that the estimate. An estimate can be used also in deciding whether to buy, hold or sell the security. For example, if the estimate exceeds the price at which the security is being offered for sale, a "buy" recommendation is appropriate.

A computed estimate can represent the value of other complex securities, e.g. a debt or insurance instrument.

Further to estimating the value of a complex security, sampling at points of a low-discrepancy deterministic sequence can be used in estimating value at risk in portfolio structuring.

Brief Description Of The Drawing

Fig. 1 is a schematic of a programmed computer system in accordance with a preferred embodiment of the invention.

Fig. 2 is a graphic representation of performance data obtained in computer trial runs with an exemplary embodiment of the invention as compared with two Monte Carlo computations.

Fig. 3 is a flow diagram of a preferred embodiment of a method in accordance with the invention.

Fig. 4 is a flow diagram of a preferred further embodiment of the invention, including the use of multiple processors.

Fig. 5 is a flow diagram of initially selling a complex security in accordance with a preferred embodiment of the invention.

Fig. 6 is a flow diagram of buying/holding/selling a complex security in accordance with a preferred further embodiment of the invention.

Fig. 7 is a flow diagram of determining value at risk (VAR) for portfolio structuring in accordance with a preferred embodiment of the invention.

Fig. 8 is a flow diagram of an example of portfolio structuring in accordance with a preferred embodiment of the invention.

Further included is an Appendix with two computer algorithms in "C" source language, respectively for computing Sobol points and Halton points. For a description of C, see B.W. Kernighan et al., The Programming Language C, Prentice-Hall, 1978.

Detailed Description Of Preferred Embodiments

Fig. 1 shows a stored-program computer 11 connected to input means 12, e.g., a keyboard, for entering financial securities data, and connected to output means 13, e.g., a visual display device, for displaying an estimated value of the financial security. The computer 11 includes a working memory 111 (M), a low-discrepancy deterministic point generator 112 (P), an integrand evaluator 113 (E), and an integrand-value combiner 114 (C).

In estimating the value of a multi-dimensional integral in financial securities valuation, a multivariate integrand

is sampled at points corresponding to a low-discrepancy deterministic sequence of points in the multivariate unit cube as defined below. If the multivariate unit cube is also the domain of integration, the points of the low-

5   discrepancy deterministic sequence can be used as sample points directly. In the case of a more general domain of integration, sample points correspond to points of a low-discrepancy deterministic sequence in the multivariate unit cube via a suitable transformation or mapping.

10   When a sufficiently large number of sampled values has been computed, an approximation of the integral is obtained by suitably combining the computed values, e.g., by averaging or weighted averaging.

In the d-dimensional unit cube $D = [0,1]^d$, a low-

15   discrepancy deterministic sequence $z_1, z_2, \ldots$ of points in D can be characterized as follows:

For a point $t = [t_1, \ldots, t_d]$ in D, define

$$[0, t) = [0, t_1) \times \ldots \times [0, t_d),$$

where $[0, t_i)$ denotes a left-closed, right-open interval,

20   and denote with $\chi_{[0, t)}(.)$ the characteristic or indicator function of $[0, t)$. For points $z_1, \ldots, z_n$ in D, define

$$R_n(t; z_1, \ldots, z_n) = (\textstyle\sum_{k=1}^{n} \chi_{[0, t)}(z_k))/n - t_1 t_2 \ldots t_d,$$

and define the discrepancy of $z_1, \ldots, z_n$ as the L_∞-norm of the function $R_n(.; z_1, \ldots, z_n)$, i.e.,

25     $$\|R_n(.; z_1, \ldots, z_n)\|_\infty = \sup_{t \in D}|R_n(t; z_1, \ldots, z_n)|.$$

The sequence $z_1, z_2, \ldots$ is said to be a low-discrepancy deterministic sequence provided

$$\|R_n(.; z_1, \ldots, z_n)\|_\infty = O((\log n)^d/n).$$

-6-

Low-discrepancy deterministic sequences are described in the published literature; see, e.g., H. Niederreiter, "Random Number Generation and Quasi-Monte Carlo Methods", $\underline{\text{CBMS-NSF}}$, 63, SIAM, Philadelphia, 1992 and S. Paskov,

5   "Average Case Complexity of Multivariate Integration for Smooth Functions", $\underline{\text{Journal of Complexity}}$, Vol. 9 (1993), pp. 291-312. Well-known examples of low-discrepancy deterministic sequences are the so-called Hammersley points, Halton points, Sobol points, and hyperbolic-cross points.

10   For illustration, in the case of Sobol points in a single dimension (d=1), a constructive definition may be given as follows: Choose a primitive polynomial

$$P(x) = x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + 1$$

(whose coefficients a_i are either 0 or 1) and define so-

15   called direction numbers v_i, $i > n$ by the following recurrence formula:

$$v_i = a_1 v_{i-1} \oplus a_2 v_{i-2} \oplus \ldots \oplus a_{n-1} v_{i-n+1} \oplus v_{i-n} \oplus (v_{i-n}/2^n),$$

where $\oplus$ denotes a bit-by-bit "exclusive or" operation.

The initial numbers $v_1 = m_1/2, \ldots, v_n = m_n/2^n$ can be

20   chosen freely provided m_i is odd and $0 < m_i < 2^i$ for $i = 1, 2, \ldots, n$.

Using the direction numbers v_i so defined, now define the one-dimensional Sobol sequence $x_1, x_2, \ldots$ by

$$x_k = b_1 v_1 \oplus b_2 v_2 \oplus \ldots \oplus b_w v_w, \quad k \geq 0$$

25   where $k = \sum_{i=0}^{\lceil \log k \rceil} b_i 2^i$ is the binary representation of k.

For higher dimensions (d > 1), the first d primitive polynomials $P_1, P_2, \ldots, P_d$ are used. If $\{x_k^i\}_{k=1}^{\infty}$ denotes the one-dimensional Sobol sequence generated by the

polynomial P_i, the d-dimensional Sobol points are defined as $x_k = (x_k^1, x_k^2, \ldots, x_k^d)$.

While this definition can be implemented as a computer algorithm directly, faster techniques are known which

5 produce these points in a "shuffled" or permuted sequence. In particular, this applies to the computer algorithm given in the Appendix.

For specificity in the following, a preferred embodiment of the invention is described as applied to a

10 collateralized mortgage obligation known as CMO FN, 89-23. This has thirty-year maturity and consists of the following tranches:

PAC tranches 23-A, 23-B, 23-C, 23-D, 23-E

supporting tranches 23-G, 23-H, 23-J

15 residual tranche 23-R

accrual tranche 23-Z

The monthly cash flow is divided and distributed according to pre-specified rules which are included in a formal prospectus. Some of the basic rules may be stated as

20 follows:

First from the monthly cash flow, the coupon is paid to the tranches. The remaining amount, called Principal Distribution Amount, is used for repayment of the principal. Prior to a fixed date in the future, the Principal Amount

25 will be allocated sequentially to the tranches 23-G, 23-H, 23-J and 23-Z. After that date, the Principal Distribution Amount will be allocated sequentially to the tranches 23-A, 23-B, 23-C, 23-D and 23-E according to a planned schedule. Any excess amount of the Principal Distribution Amount over

the planned schedule will be allocated sequentially to the tranches 23-G, 23-H, 23-J and 23-Z. A distribution of principal of the tranche 23-R will be made only after all other tranches have been retired.

In deriving an estimate for the present value of the security at the time of issue, the following notation is used below:

C   - the monthly payment on the underlying mortgage pool;

i_k - the projected interest rate in month k, k = 1, 2, ..., 360;

w_k - the percentage of mortgages prepaying in month k;

$a_{360-k+1}$ - the remaining annuity after month k.

A remaining annuity a_k can be expressed as

$$a_k = 1 + v_0 + \ldots + v_0^{k-1}, \quad k = 1, 2, \ldots, 360,$$

with $v_0 = 1/(1+i_0)$, where i_0 is the current monthly interest rate. Thus, after k months, the remaining amount of principal borrowed is $C \cdot a_k$.

It is assumed that the interest rate i_k can be expressed as

$$i_k = K_0 \, \exp(\xi_k) \, i_{t-1},$$

where exp(.) denotes exponentiation and where $\xi_1, \xi_2, \ldots, \xi_{360}$ are independent, normally distributed random variables with mean 0 and variance σ^2, and K_0 is a given constant. For the present example, $\sigma^2 = 0.0004$ is chosen.

It is assumed further that w_k as a function of i_k can be computed as

$$w_k = K_1 + K_2 \arctan(K_3 i_k + K_4),$$

where K_1, K_2, K_3, K_4 are given constants.

-9-

Under these assumptions. the cash flow in month k, k = 1, 2, ..., 360 is

$$C(1-w_1(\xi_1))\ldots(1-w_{k-1}(\xi_1,\ldots,\xi_{k-1}))(1-w_k(\xi_1,\ldots,\xi_k)$$
$$+ w_k(\xi_1,\ldots,\xi_k)a_{360-k+1}),$$

5   where

$$w_k(\xi_1,\ldots,\xi_k) = K_1 + K_2 \arctan(K_3 K_0^k i_0 \exp(\xi_1+\ldots+\xi_k)+K_4).$$

This cash flow is distributed according to the rules of FN, 89-23. Then, the cash flow for each of the tranches is multiplied by the discount factor

10    $$v_1(\xi_1)\ldots v_k(\xi_1,\ldots,\xi_k), \quad \text{with}$$

$$v_j(\xi_1,\ldots,\xi_j) = 1/(1 + K_0^j i_0 \exp(\xi_1+\ldots+\xi_j)), \quad j=1,2,\ldots,360,$$

to find the present value for month k. Summing of the present values for every month gives the present value PV_T, for each tranche T.

15   The expected or estimated value,

$$E(PV_T) = E(PV_T(\xi_1,\ldots,\xi_{360})),$$

upon a change of variables is represented by

$$E(PV_T) = \int_D PV_T(y_1(x_1),\ldots,y_{360}(x_{360}))dx_1\ldots dx_{360},$$

where

20    $$x_i = (2\pi\sigma)^{-\frac{1}{2}}\int_{-\infty}^{y_i} \exp(-t^2/(2\sigma))dt.$$

Thus, for each tranche T, a 360-variate integrand has to be integrated over the 360-dimensional unit cube, .

After generating a point

$$(x_1, x_2, \ldots, x_{360})$$

25   of a low-discrepancy deterministic sequence in the unit cube, the point

$$(y_1, y_2, \ldots, y_{360})$$

/1

is computed by finding the value of the inverse normal cumulative distribution function at each x_k. Then, for each tranche T, the function value

$$PV_T(y_1, \ y_2, \ \ldots, \ y_{360})$$

5   is computed. These are the function values used in numerical integration.

Fig. 2 shows results from trial runs for CMO FN, 89-23 with a preferred method using Sobol points generated by the corresponding computer algorithm given in the Appendix, as

10   compared with Monte Carlo integration. Two Monte Carlo computations were carried out, with different "seeds" or starting values of a congruential pseudo-random number generator known as RAN2; see W. Press et al., <u>Numerical</u> <u>Recipes in C</u>, Cambridge University Press, 1992. It is

15   apparent that the preferred method reaches a steady value more rapidly. In this and other trial runs, with typical requirements of precision and confidence, a speed-up by a factor of 3 to 5 was realized as compared with Monte Carlo integration. Much greater speed-up can be expected when

20   higher precision or/and higher levels of confidence are sought.

In a further trial run with CMO FN, 89-23, instead of Sobol points, Halton points were used as generated by the corresponding computer algorithm given in the Appendix. It

25   is felt that Sobol points may be preferred over Halton points for integrals of high dimension. However, this preference may not apply in the case of lower-dimensional integrals, e.g., with dimension up to 5 or so.

A computation as described may be terminated after a
predetermined number of function evaluations.
Alternatively, e.g., after every function evaluation or
after a predetermined incremental number of function

5   evaluations, a current approximation may be compared with
one or several preceding approximations, for termination
once a suitable condition depending on the difference
between approximations is met. Such termination criteria
may be called automatic. Automatic termination is

10   particularly reliable where a sequence of approximations
settles down smoothly; see, e.g., the curve in Fig. 2
corresponding to Sobol points.

Advantageously in computing function values, a cluster
or network of multiple parallel processors or workstations

15   can be used. This may involve a master or host processor
providing points of a low-discrepancy sequence to slave
processors and combining function values returned by the
slave processors into an approximate value for the integral.
Thus, the computation can be speeded up in proportion to the

20   number of processors used.

Advantageous further, in combination with numerical
integration as described above, is the use of error
reduction techniques analogous to variance reduction in
Monte Carlo integration as described, e.g., by M. Kalos et

25   al., Monte Carlo Methods, John Wiley & Sons, 1986. This may
involve a change of variables or/and variation reduction,
for example.

The specific embodiment described above, involving
valuation of a collateralized mortgage obligation, can be

interpreted as follows: The definite integral represents
the expected value of the sum of a series of payments on the
obligation, namely the regular payments of mortgage
principal and interest, plus any mortgage prepayments. The
5   likelihood of prepayment depends in part on the prevailing
level of interest for mortgage refunding.

For each payment, the value of the payment is weighted
by the probability of all interest rate scenarios up to the
time of the payment. Thus, for example, the amount of next
10   month's payment will be valued based on the probability of
all possible interest rates likely to prevail next month,
and the impact of each of those possibilities on the
likelihood of prepayment of principal. For the month after,
valuation takes into account the interest rate scenarios of
15   both months, their probabilities of occurrence, and the
effect of interest rates on prepayment. In addition to next
month's valuation, the month-after estimate will include an
additional term corresponding to the month-after payment,
represented by an integral over the possibilities of the
20   two-month process. Further monthly payments, i.e. third,
fourth et seq., are represented analogously. With each
additional integral term representing a scheduled or
unscheduled payment amount, the over-all value of the CMO is
represented by the sum of the terms.

25   The value of the CMO can be decomposed further, e.g.,
by valuing the interest payment and the principal payment
separately in each period to represent an "interest-only
tranche" and a "principal-only tranche". Thus, the value of
the CMO can be represented by the sum of a yet-greater

number of integrals. Conveniently, the value of each of the individual integrals in the sum can be estimated as described above, using a low-discrepancy deterministic sequence of sample points.

5   A sum of payment integrals can also be used to represent the value of a portfolio of two or more CMOs, giving the same result as a single integral over the sum of the payments. The range of possibilities over which the integration is effected is the same in both cases.

10   A CMO portfolio can be augmented in other ways, e.g., by the inclusion of one or several bonds. For valuation purposes, a high-grade bond differs from a CMO in that prepayment will take place for the entire issue or not at all. Again, valuation depends on the fluctuation of
15   interest rates, taking into account all possible future interest rate scenarios. This applies whether there is one payment (from a single CMO), several payments (from several CMOs), or payments from a combination of CMOs and bonds.

       Other than CMOs and bonds, a complex security portfolio
20   may include securities such as single-premium deferred annuities (SPDA). An SPDA is a contract sold by an insurance company to an annuitant investor. Under the contract, the insurance company takes money from the investor, invests this money in a portfolio of various
25   investments, and credits a portion of the investment proceeds to the investor's account. The SPDA is an asset of the insurance company. There is an initial investment by the company in the cost of producing the contract. There is a predictable return for the company on the investment

portfolio, of interest and principal payments by the various borrowers less the contractual crediting to the annuitant's account. The unpredictable factor in the valuation of the portfolio is again random movements of interest rates, as

5   these influence prepayment of investment assets as well as unexpected withdrawals and cash demands by the annuitant. The same methodology of valuation can be used after inclusion of the SPDA asset in the complex security.

Valuation, by the insurance company, of a portfolio

10   including an SPDA contract involves estimating the value of the contract itself. In pricing a new contract, its value is estimated in advance of issue by making assumptions about the kinds of assets that will be in the portfolio along with the SPDA. The pricing of the SPDA before issue and its

15   valuation after issue are equivalent to the calculation presented above for a CMO.

A life insurance policy, or any other kind of insurance contract differs from the above in that, in addition to the interest rate variable, there is an additional stochastic

20   variable representing health risk, mortality risk or fire risk, for example. Under inclusion of this variable, the same techniques as described above can be used for valuation.

Many other types of securities are amenable to

25   corresponding valuation and estimation, where each period's prospective payment is valued depending upon all previous values of a stochastic variable. For example, in the case of an interest rate option, the stochastic variable is interest rate and the payment is the value of a bond. In

the case of a stock option, the variable is the price of the stock. In the case of a currency option, the variable is the relative value of two currencies.

As an example of an application of sampling of a function other than an integrand, the following concerns value at risk (VAR). If

$$h = h(\xi_1, \ldots, \xi_d)$$

denotes the potential loss of a portfolio whose value depends on d market factors represented by stochastic variables $\xi_1, \ldots, \xi_d$, and if F_h denotes the cumulative distribution function of h, i.e.

$$F_h(x) = \text{Prob}\ \{h < x\},$$

then, for a specified percentile or quantile p,

$$VAR = x_p = F_h^{-1}(p).$$

The value of p represents a level of confidence, with p=1 representing certainty. For practical purposes, a choice of p=0.98 is typical, indicating that the loss may exceed VAR with a probability of 2%.

In accordance with a preferred embodiment of the invention, VAR can be estimated by a technique as illustrated by Fig. 7, using sampling of the potential loss function at m points of a low-discrepancy deterministic sequence. The potential loss h_k is evaluated for each point of the sequence, resulting in a list $h_1, \ldots, h_m$. The number of points, can be pre-set, e.g. at m=1000. The list is sorted, and the cumulative distribution function is determined. The cumulative distribution fucntion is a monotonically increasing function whose domain {x} includes the h-values and whose range {$F_h(x)$} is in the interval from

-16-

0 to 1. The desired estimate of VAR is determined as that value of x for which $F_h(x)=p$. (It is understood that, wherever exact equality cannot be realized in calculation, a value will be approximated or interpolated.)

5   An implementation of the technique was applied to a 34-dimensional problem, namely a portfolio of 34 at-the-money currency and equity options having maturity of one year, with $10,000 allocated to each of the options, yielding a total portfolio of $340,000. Volatility and correlation

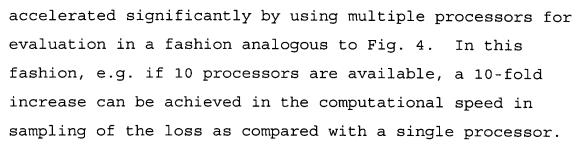10   data were taken from J.P.Morgan/Reuters RiskMetrics data sets.

    The technique was applied further to a 360-dimensional problem, namely the CMO portfolio consisting of ten tranches as described above, with $100,000 allocated to each tranche,

15   yielding a total portfolio or $1,000,000.

    For the options portfolio, 1,000 points of the low-discrepancy deterministic generalized Faure sequence resulted in VAR being determined to within 1% of a reference determination using 100,000 points. For the CMO portfolio,

20   a determination of VAR using 1,500 points of the same sequence also was within 1% of a reference value, obtained from 150,000 points. These results are contrasted with prior-art Monte-Carlo computations where, for the same numbers of samples (1,000 and 1,500, respectively), VAR

25   could be estimated only to within 10% of the respective reference values, i.e. with 10-fold error as compared with the low-discrepancy deterministic sequence determinations.

    In the technique as described, computational effort is greatest in evaluating the h_k's, and the computation can be

18

accelerated significantly by using multiple processors for evaluation in a fashion analogous to Fig. 4. In this fashion, e.g. if 10 processors are available, a 10-fold increase can be achieved in the computational speed in

5   sampling of the loss as compared with a single processor.

As illustrated by Fig. 8, a computed value of VAR can be compared with a target value which is considered acceptable. If VAR meets the target level, the portfolio may remain unchanged; if VAR is considerably less than the

10   target level, the portfolio may be restructured for assumption of greater risk and potentially increased return; if VAR exceeds the target level, restructuring for reduced risk is indicated. Restructuring may be performed immediately or over time, e.g. by changing investment

15   policies affecting future transactions.